Procedures for the Interpretation of Nuclear DNA Typing Results from the Yfiler™ PCR Amplification Kit

1 Scope

These procedures apply to DNA personnel who verify and interpret nuclear DNA typing results obtained from the Yfiler™ PCR Amplification Kit using the GeneMapper™ *ID-X* (GMIDX) DNA typing software for forensic comparison purposes and the Y-Chromosome Haplotype Database for statistical analysis.

2 Background

Y-STR examinations are generally conducted for lineage comparisons, in support of Missing Persons and Intelligence cases, or on samples that have a high ratio of female to male DNA.

Y-STR profiles are considered haplotypes so all conclusions regarding comparisons apply to the person of interest (POI) and the paternal lineage of the POI. In addition, unrelated individuals may exhibit the same Y-STR typing results.

3 Equipment/Materials/Reagents

GeneMapper™ ID-X software (Applied Biosystems, version 1.5 or higher)

Y-Chromosome Haplotype Reference Database, current release, www.yhrd.org (Institute of Legal Medicine, Charite – University Medicine Berlin)

4 Standards and Controls

Raw data for the electrophoretic runs of samples or controls displaying no typing results must be reviewed for the presence of a primer peak. If no primer peak is observed, the sample must be reinjected or reprepared to verify that amplicon was added to the capillary electrophoresis (CE) plate.

4.1 Verification of GeneScan®-500 (GS-500) Internal Size Standard (ISS)

4.1.1 For Yfiler™ data, which is sized using the Local Southern Method, verify that the 75, 100, 139, 150, 160, 200, 300, 340, and 350 base pair (bp) fragments are captured and have been assigned the correct size values for each sample, control, and allelic ladder. In addition, the 400 bp fragment must be captured and properly assigned for any sample containing an allele ≥340 bp to be used for interpretation. Due to the temperature sensitivity of the 250 bp fragment's sequence-based conformation, this fragment is not used for sizing purposes.

4.1.2 If all of the GS-500 fragments for a given injection do not meet these specifications, a different injection of the sample that displays the correct size values for all of the GS-500 fragments must be used for interpretation of the entire DNA profile, which may require that the sample be reprocessed.

4.2 Verification of Allelic Ladders

4.2.1 Any allelic ladder used for genotyping must: 1) exhibit the correct allele designations (see Table 1) and 2) yield the correct typing results when used to genotype the positive amplification control.

4.2.2 If any sample(s) requires reinjection the appropriate ladder must be included in the reinjection set.

Locus[1]	Known Size Range (bp)[2]	Alleles Present in Ladder[3]	Color
DYS456	104-123	13-18	Blue
DYS389I	142-163	10-15	Blue
DYS390	191-227	18-27	Blue
DYS389II	252-293	24-34	Blue
DYS458	130-155	14-20	Green
DYS19	175-210	10-19	Green
DYS385a/b[4]	242-317	7-25	Green
DYS393	100-131	8-16	Yellow
DYS391	150-175	7-13	Yellow
DYS439	197-225	8-15	Yellow
DYS635	246-270	20-26	Yellow
DYS392	290-325	7-18	Yellow
Y GATA H4	121-142	8-13	Red
DYS437	182-197	13-17	Red
DYS438	223-248	8-13	Red
DYS448	280-324	17-24	Red

Table 1 – Yfiler™ Allelic Ladder Specifications

[1] Yfiler™ haplotypes consist of tetranucleotide repeats with the exception of the trinucleotide repeat DYS392, pentanucleotide repeat DYS438, and hexanucleotide repeat DYS448.
[2] Sizes in base pairs are approximate due to electrophoretic variation and are based on plus-A addition. These sizes are published in the *AmpFlSTR® Yfiler™ PCR Amplification Kit User's Manual* (Applied Biosystems, Inc.).
[3] Ranges of alleles (i.e., 13-18) include only integers (i.e., 13, 14, 15, …, 18).
[4] DYS385 is a multi-copy (a/b) locus and may exhibit a single allelic peak or two different allelic peaks.

4.3 Positive Amplification Control (i.e., 007)

4.3.1 One positive control must be processed for amplification in parallel with each set or batch of evidentiary samples.

4.3.1.1 If any sample(s) requires repreparation, the positive control and the appropriate ladder must also be reprepared.

4.3.2 If a set of samples has multiple injections of the positive control, at least one injection must exhibit all the expected allelic peaks ≥50 relative fluorescence units (RFUs) and must not exhibit any extraneous allelic peaks. A positive control with a non-allelic peak(s) (e.g.,, stutter, spike, pull-up) may be interpreted. See Table 2 for the expected positive control typing results using Yfiler™:

Locus	007 Control
DYS456	15
DYS389I	13
DYS390	24
DYS389II	29
DYS458	17
DYS19	15
DYS385a/b	11,14
DYS393	13
DYS391	11
DYS439	12
DYS635	24
DYS392	13
Y GATA H4	13
DYS437	15
DYS438	12
DYS448	19

Table 2 – Expected Yfiler™ STR Typing Results for the 007 Positive Control

4.3.2.1 Refer to the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233) for guidance if the positive amplification control does not exhibit the expected results.

4.4 Negative Amplification Control

Refer to the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233) for guidance on the evaluation of the negative amplification control.

4.5 Extraction Control (i.e., reagent blank)

Refer to the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233) for guidance on the evaluation of the extraction control.

5 Procedures

5.1 DNA Profile Determination

5.1.1 Computer Assisted Allele Designations

The GMIDX software, using the analysis method settings represented in Appendix A, analyzes the data generated by the CE instruments and generates electropherogram data to be evaluated and interpreted. A pink box surrounding a data point label indicates that the software has identified a data point as an artifact. The GMIDX software uses the terms "spike" and "OMR" (Outside Marker Range) to represent a variety of DNA artifacts. Peak labels may be edited according to this procedure. Peaks interpreted as non-allelic may be deleted within GMIDX and will appear on the electropherogram with a single strikeout.

5.1.1.1 Identification of Peaks of Non-Genetic (Non-Allelic) Origin

Before the Y-STR typing results from a sample can be used for comparison purposes, it is necessary to identify any non-genetic peaks that do not represent allelic Y-STRs. These non-genetic peaks may be undesired PCR products (e.g., stutter, minus-A, and non-specific product), analytical artifacts (e.g., spikes and raised baseline), instrumental limitations (e.g., matrix failure), or be introduced into the process (e.g., dissociated primer dye and non-specific peaks).[5]

Reproducible non-genetic peaks (e.g., stutter, non-template dependent nucleotide addition, dissociated dye, matrix failure, non-specific product) may be interpreted.[6] Non-reproducible non-genetic peaks (e.g., spikes and raised baseline) must be evaluated as specified.

Refer to the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233) for information regarding excessive DNA template and off-scale samples, raised sample baseline, non-template-dependent nucleotide addition, matrix failure resulting in pull-up, spikes, and dissociated primer dyes.

[5] The GMIDX software applies two levels of filtering to the sized Y-STR allelic data. The first filter is global and removes labels from peaks that are less than 2% of the peak height of the largest allele present at each Y-STR locus. The second filter removes labels from peaks at any locus that meet the FBI-defined sizing and relative peak height criteria for stutter and/or minus-A.

[6] For purposes of interpreting DNA typing results, a peak need only be identified as being of non-genetic origin.

5.1.1.1.1 Stutter

The kit-specific stutter percentage guidelines provided in Table 3 are estimates (Average + 3 SD) of the maximum expected stutter values at each locus in the Yfiler™ amplification kit. These values are expressed as a percentage relative to the source allelic peak height (i.e., % Stutter). Stutter peaks have been observed up to +/- 3 repeat units away from the source allelic peaks in both the absence and presence of intermediate stutter peaks. If such atypical stutter peaks are due to excessive amounts of template DNA, the sample may be reamplified with less template DNA or reinjected for less time.

Locus	-1 Repeat Unit	+1 Repeat Unit
DYS456	16	6
DYS389I	10	6
DYS390	13	6
DYS389II	18	6
DYS458	14	6
DYS19	13	6
DYS385a/b	17	6
DYS393	14	6
DYS391	10	6
DYS439	10	6
DYS635	13	6
DYS392	15	8
Y GATA H4	11	6
DYS437	8	6
DYS438	5	8
DYS448	6	5

Table 3 – Maximum Expected Yfiler™ STR Loci Stutter Percentages

5.1.1.1.2 Non-Specific Peaks

a. An N-2 peak is frequently observed at the DYS19 locus as an off-ladder (OL) allele. This artifact is generally reproducible.

b. A reproducible minus-A peak of the N-4 stutter peak may occur at DYS437.

5.1.1.1.3 Additional peaks of non-genetic origin are described in the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233).

5.1.1.2 Off-Ladder (OL) Alleles

If an allele fails to size within a defined allele category (e.g., a bin or a virtual bin), it must be assigned a size using the following criteria.

5.1.1.2.1 Any sample containing an OL allele may be re-injected.

5.1.1.2.2 An OL allele may be a microvariant that sizes between two ladder alleles. For example, if an OL allele occurs between the 12 and 13 ladder alleles and is approximately 1 bp larger than the 12 allele, it is designated as 12.1; 2 bp larger is designated 12.2; and so on.

5.1.1.2.3 If an OL allele does not fall within the size range of any locus-specific ladder, which includes the flanking virtual bins, it should be associated with one of the two loci between which it falls.

a. For single-source samples, if the OL allele is flanked by a locus with one peak and a locus with no peaks (with the exception of DYS385 a/b), the OL allele is assigned to the latter locus.

b. Generally, if both loci between which an OL allele falls each display a single allele, the OL allele may be assigned to the locus closest in size.[7]

c. If determination of the locus assignment is not possible, both loci that flank the OL allele must be deemed inconclusive for matching/statistical purposes.

d. If the OL allele is smaller in size than the smallest respective virtual bin, or larger in size than the largest respective virtual bin, the number of repeats in the allele should be estimated.

e. When loci are closely spaced on the x-axis of an electropherogram, an above or below OL allele may be observed within the size range of a flanking locus.

5.2 Contamination Assessment

Refer to the appropriate DNA procedure for the interpretation of nuclear DNA typing results (i.e., DNA 233) for guidance on the evaluation of contamination in samples or controls.

5.3 Application of Peak Height Thresholds to Allelic Peaks

5.3.1 The Analytical Threshold (AT) is 50 RFU. At all loci other than DYS385a/b, any allelic peak detected at or above the AT may be used for matching/statistical purposes.

[7] To facilitate the interpretation of OL alleles, the Examiner may consult a listing of such alleles recorded at http://www.cstl nist.gov/div831/strbase/var_tab.htm.

5.3.2 The Hemizygous Interpretation Threshold (HIT) is 125 RFU and is an empirically determined parameter established specifically for the DYS385a/b locus. The HIT is used to evaluate potential allelic drop-out in a profile.

5.3.2.1 The Y-STR typing results at DYS385a/b may be used for matching/statistical purposes if two alleles are detected or if one allele is detected at or above the HIT. If only one allelic peak is detected below the HIT, the locus may be used only for exclusionary purposes. However, if Y-STR or autosomal STR typing results indicate the presence of DNA from more than one male contributor, then two alleles detected below the HIT at DYS385a/b may not be used for matching/statistical and/or exclusionary purposes.

5.3.3 If no allele is detected at a locus, a null allele may be declared at that locus if all allelic peaks at all other loci are at or above the HIT.

5.4 Interpretation of Y-STR Typing Results

To the extent possible, DNA typing results from evidentiary samples will be interpreted before the comparison to any known samples, other than those of assumed contributors.

When there are multiple amplifications and/or injections for a given sample extract, generally the one that provides the most information will be used for reporting. Data where saturation interferes with interpretation may require that an alternative amplification/injection is used.

5.4.1 Peak Height Ratio Assessment

For loci that contain more than one allelic peak, peak height ratios can be used to associate two alleles to a common source (i.e., duplication) or to establish the presence of a DNA mixture. Peak height ratio assessments are generally not used in the interpretation of knowns and items that are expected to have originated from a single source, such as bones and alternate knowns. The alternate known profile may be obtained as a single-source or as a major contributor typing result.

5.4.1.1 Peak height ratios (PHR) are calculated by dividing the peak height of the allele with the lower RFU value by the peak height of the allele with the higher RFU value, and then expressed as a percentage.

5.4.2 Determination of the Number of Contributors to Y-STR Typing Results

5.4.2.1 A profile is generally considered to have originated from a single male individual if one allele (other than DYS385a/b) is present at all loci for which typing results were obtained.

5.4.2.2 A profile is generally considered to have originated from more than one male individual if two or more alleles are present at two or more loci, other than DYS385a/b. The classification of any DNA profile as a mixture must be based on an evaluation of the DNA profile in its entirety.

5.4.2.3 Peaks that exceed the expected stutter percentages must be evaluated considering:

- A peak significantly above the stutter percentage is more likely to be allelic.
- A peak at a small (<200 bp) locus where possible minor contributor types are expected has more potential to be allelic.
- A peak in an additive stutter position, which exceeds the negative stutter percentage but not the combined positive and negative stutter percentages, may be considered stutter.
- Other apparent peaks below the AT suggest that the peak is potentially allelic.
- If the sample is a reference sample and expected to be single source, then these peaks can confidently be called stutter if there is no other evidence of contamination.

5.4.2.3.1 For apparent single source samples, a peak in a stutter position that exceeds the expected stutter percentage may be interpreted as a stutter peak for purposes of determining the number of contributors to the sample. Generally, this interpretation is limited to a single instance unless the peaks are at a large (>200 bp) locus and there are no other indications of a mixture (e.g., peak height imbalance, apparent peaks <AT).

5.4.2.3.2 For mixed samples, peaks that exceed the expected stutter percentage are generally considered allelic for purposes of determining the number of contributors to the sample.

5.4.2.4 Duplications

- If two alleles are present at any locus other than DYS385a/b, a duplication event may have occurred. To be declared a duplication, the alleles should have a PHR ≥ 40%.
- Because a duplicated allele is typically one repeat unit larger or smaller than the other allele, the presence of two alleles at a given locus that differ in size by more than one repeat unit is generally indicative of a mixed profile.
- The proximity of certain loci along the Y-chromosome allows for the simultaneous duplication of alleles at multiple loci. Generally, loci that are less than 1 Mb apart could potentially be duplicated together.[8] See Appendix B for the relative positioning and distance (i.e., Mb) of the Yfiler™ loci on the Y-chromosome.
- A profile in which three allelic peaks are observed at a single locus, but in which no other typing results indicate the presence of a mixture, may be concluded to be a single-source profile possessing a tri-allelic locus.
- The presence of more than one allele at DYS438 and/or Y GATA H4 is generally indicative of a mixed profile.

5.4.2.5 The number of contributors to a mixture should be based upon greatest number of alleles detected per locus, and, because of the potential for duplication, should generally be

[8] A Y-STR haplotype exhibiting duplication at loci DYS437, DYS439 and DYS389I/II has been observed (Butler, JM, Decker, AE, Kline, MC, Vallone, PM. 'Chromosomal Duplications Along the Y-Chromosome and Their Potential Impact on Y-STR Interpretation'. J For Sci (2005), 50(4), 853-859).

observed at two or more loci (other than DYS385a/b). This is the initial estimate of the number of contributors to the sample.

5.4.2.6 Using the loci with the largest number of alleles, assess the ratio of contributors. Evaluate peak height imbalance and account for allele sharing to determine if the number of contributors should be increased.

5.4.2.7 Apply the general pattern of number of contributors and mixture ratio across the profile to determine if other loci are consistent with this pattern or if the number of contributors should be increased or decreased by one. Loci with more alleles will be the most informative for this assessment. Additionally, apparent peaks <AT may also be considered, especially for low level samples.

5.4.3 Deduced Single-Source Profiles Determined by Separation of Expected Typing Results

5.4.3.1 For mixed typing results, when the presence of an individual's DNA in the sample can be reasonably expected, the Y-STR typing results from the assumed contributor should be separated from the other mixture results to facilitate identification of the foreign alleles.

5.4.3.2 If sharing of alleles among the conditional male known profile and an additional male is suspected at a locus, no separation of each individual's alleles is possible at that locus.

5.4.3.3 This approach can also be used when another known male is expected to have contributed biological material to the mixed profile (e.g., consensual male). If more than one known male contributor is expected, each individual's alleles should be subtracted from the profile.

5.4.3.4 This approach can also be applied to evidentiary items from which DNA is isolated by means of a differential extraction. In such situations, the single-source or major contributor typing results from one fraction may be used as a conditional known profile(s) applied to the complementary fraction.

5.4.3.5 Mixtures comprised of three or more individuals generally may only be used for exclusionary purposes. If an apparent distinguishable mixture remains after subtraction of the assumed contributor, consult with the Technical Leader (TL) for additional guidance.

5.4.4 Deduced Single-Source Profiles Determined from Distinguishable Mixtures

5.4.4.1 A distinguishable mixture is a Y-STR typing result from a sample for which alleles can be attributed to individual major/minor male donors. In order to determine major/minor contributors to a mixture, every locus must be distinguishable except for DYS385a/b.

5.4.4.1.1 When a mixed profile contains at least one allele above the AT at every locus (i.e., full profile), the following should be applied:

a. The PHR of alleles at all loci that exhibit two alleles (with the exception of DYS385a/b) must be ≤ 55% in order to assign a major contributor. When all loci meet this requirement, the allele displaying the greater peak height at each locus may be attributed to the major contributor.

b. If only one allele is detected at a given locus, that allele may be attributed to the major contributor.

5.4.4.1.2 When a mixed profile fails to produce alleles above the AT at one or more loci (i.e., partial profile), the following should be applied:

a. The PHR at all loci that exhibit two alleles (with the exception of DYS385a/b) must be ≤ 40% in order to assign a major contributor. When all loci meet this requirement, the allele displaying the greater peak height at each locus may be attributed to the major contributor.

b. If only one allele is detected at a given locus, that allele may be attributed to the major contributor.

5.4.4.2 Mixtures comprised of three or more individuals generally may only be used for exclusionary purposes. If a major contributor can be discerned from a mixture comprised of three or more individuals, consult the Technical Leader for guidance.

5.4.4.3 Due to the possibility that the minor contributor's alleles may be either shared by the major contributor (and thus masked) or potentially not detectable (i.e., <50 RFU), determination of the minor contributor profile may be possible at only some loci.

5.4.4.4 Generally, a multi-locus, mixed sample that contains one or more true minor contributors can be expected to display at least one allelic peak in a non-stutter position.

5.4.5 Interpretation of Y-STR Typing Results for Indistinguishable Mixtures

An indistinguishable mixture is a Y-STR typing result from a sample for which alleles cannot be attributed to individual donors. Indistinguishable mixtures may be used for exclusionary purposes only.

6 Reporting Y-STR Results and Conclusions

The results and/or conclusions for specimens subjected to DNA analysis will generally be reported in narrative form with tables for statistical information, when applicable.

6.1 Only single-source or deduced single-source Y-STR profiles may be used for matching or statistical purposes.

6.2 Indistinguishable mixtures are not suitable for matching purposes and may be used for exclusionary purposes only.

6.3 The number of contributors must be determined for mixtures used for comparison purposes.

6.4 An exclusion is declared when two single source or deduced single source profiles are different at one or more loci for direct comparisons, or at two or more loci for paternal lineage comparisons.[9] Differences that could be explained by allelic dropout cannot be the basis for an exclusion.

6.4.1 For a mixture comparison, an exclusion is declared when the types of a known reference item cannot be included in the types present at the corresponding loci of the mixture profile, considering the number of contributors and the potential for dropout. An exclusion for a direct mixture comparison requires differences at one or more loci, while an exclusion for a lineage mixture comparison requires differences at two or more loci.

6.5 A paternal lineage comparison is declared inconclusive when profiles from putative male relatives from the same paternal lineage are found to differ at only one[10] locus, considering the potential for dropout.

6.6 Based on the structure of a Y-STR locus, a single mutation can cause differences at one or more loci when comparing results from individuals in the same paternal lineage. For example, at DYS389I & II, a single mutation may result in differences at both the DYS389I and the DYS389II loci. Similarly, a single mutation may result in an apparent duplication at DYS437 as well as an apparent null at DYS448. These differences alone cannot be the basis of a paternal lineage exclusion, and such comparisons must be reported as inconclusive.

6.7 For both direct and lineage comparisons, an inclusion is declared when profiles are found to be the same at all loci for which interpretable DNA typing results were obtained.

6.7.1 Each DNA association or inclusion must be clearly and properly qualified with either a statistic or a qualitative statement. A qualitative statement not based on a statistical calculation should be limited to situations in which the presence of an individual's DNA on an item is reasonably expected. The provenance of the sample must be established in the case record when statistics are not calculated.

6.7.1.1 The Y- Chromosome Haplotype Reference Database (YHRD) calculates 95% Upper Confidence Interval (95% UCI) values using several general United States population groups: African American, Asian, Caucasian, Hispanic, and Native American. The Eskimo Aleut population may also be calculated. For all samples, the LRs for African American, Caucasian, and Hispanic are considered for reporting. For cases submitted from Native American land, the

[9] One mismatch is not exclusionary in paternal lineage comparisons due to the potential for mutation.
[10] For lineage comparisons, a single difference between two items may occur due to (a) mutation among relatives in the same paternal lineage or (b) different paternal lineages.

Native American LR is also considered. For all cases submitted from Alaska, the Native American LR and the Eskimo Aleut LR are also considered.

6.7.1.2 The single lowest LR value across all populations considered is reported, and it is generally truncated to two significant digits for reporting. However, if this value is less than 10, it is truncated to one significant digit for reporting.

6.7.1.3 The magnitude of the LR relates to the degree of support provided by the evidence under the tested hypotheses and assumptions. A qualitative statement will be reported based on the following table:

LR	Qualitative Equivalent
1	Uninformative
2 to <100	Limited support for **Inclusion**
100 to <10,000	Moderate support for **Inclusion**
10,000 to <1,000,000	Strong support for **Inclusion**

Table 4 – Qualitative Equivalent Scale for Y-STR Likelihood Ratios

7 Calculations

7.1 Calculation of Statistics Using the Y- Chromosome Haplotype Reference Database (YHRD)

The primary database used for estimating the haplotype frequency is the Y-Chromosome Haplotype Reference Database (YHRD; https://yhrd.org), which is used to provide Likelihood Ratios from frequency estimates for African American, Caucasian, Hispanic, Native American, and Eskimo Aleut population groups, when applicable.

7.1.1 Go to the YHRD website (http://yhrd.org). Select the "Search the Database" tab. Whenever possible, for profiles where electropherograms reflect the searched profile, choose the "search using your … GeneMapper® ID/ID-X…" option to import Y-STR profiles directly from an exported GMIDX file. Alternately, select "Manually enter the haplotypes/haplotypes…" (Figure 1).

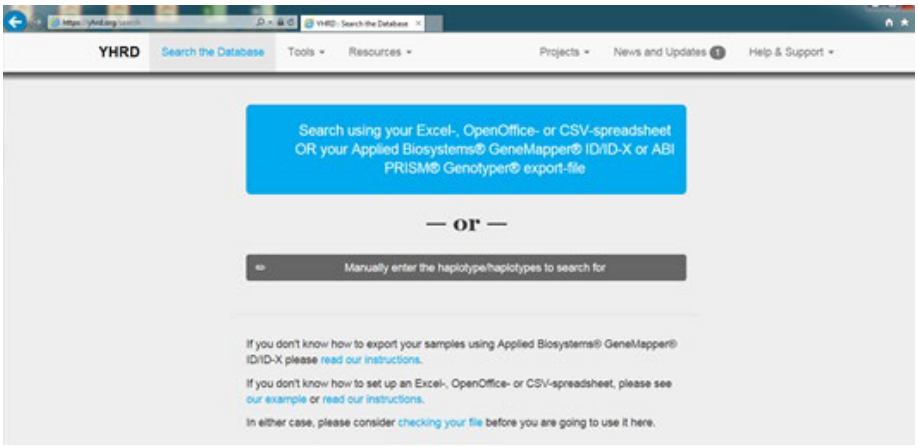

Figure 1 – YHRD Search the Database Page

7.1.2 Exported GMIDX files for STaCS entry can be used for import into YHRD. For electronically imported profiles, designate each profile to be searched, choose "Yfiler", then select "Search". (Figure 2)

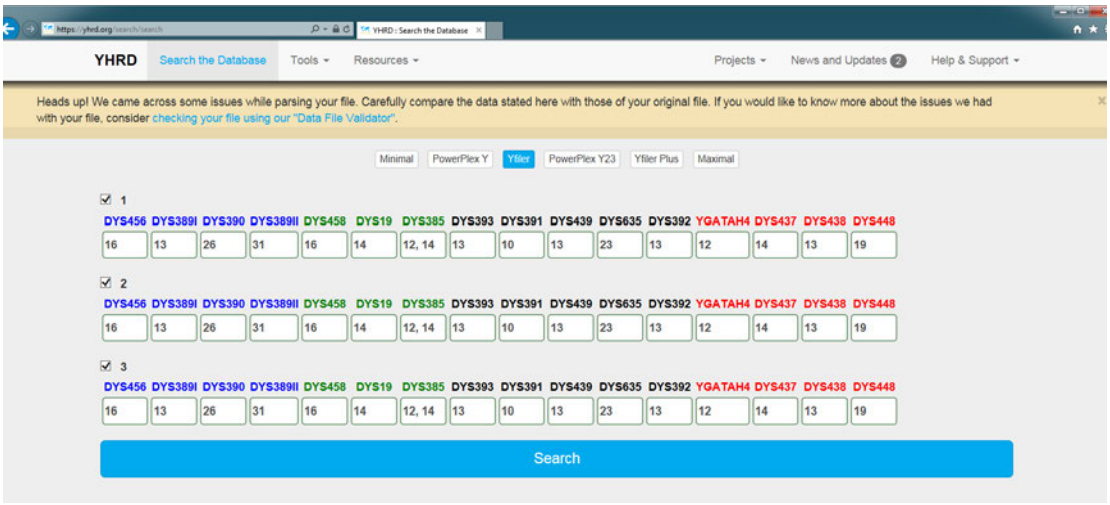

Figure 2 – YHRD Electronic Data Import

7.1.3 For manually entered profiles, choose "Yfiler" and enter the alleles for the appropriate markers, then select "Search" (Figure 3).

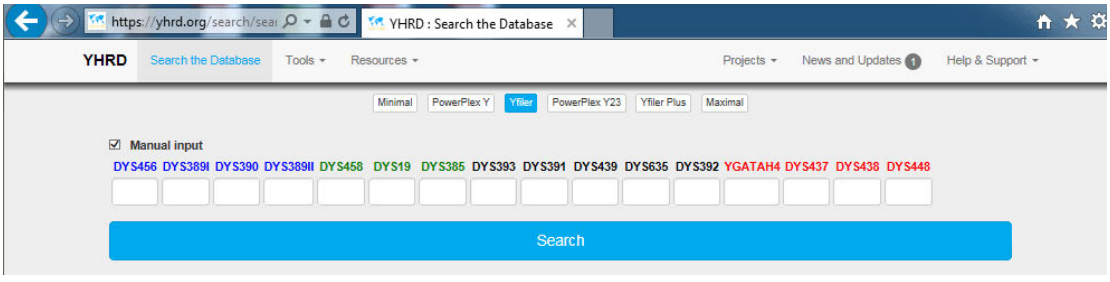

Figure 3 – YHRD Manual Data Input

7.1.4 For each sample, select "Add feature to this report", and choose "National Database (with Subpopulations, 2014 SWGDAM-compliant)". Note that YHRD will only search against the database haplotypes that contain the same or more loci than were entered.

7.1.5 The number of observations of the searched haplotype is listed for the African American, Asian, Caucasian, Hispanic, and Native American United States subpopulations.

7.1.6 If necessary, add the Eskimo Aleut subpopulation. This can be done by selecting "Add feature to this report" and choosing "Metapopulation". Change the default "Eurasian - European" metapopulation in the report to "Eskimo Aleut". Change the default two-sided 95% CI calculation to "95% Upper Confidence Interval (UCI)".

7.1.7 For instances in which the haplotype has been observed in the database (x > 0), YHRD calculates a one-sided 95% confidence interval using the Clopper and Pearson formula:[11]

$$\sum_{k=0}^{x}\binom{n}{k}p_0^k(1-p_0)^{n-k}=\alpha$$

where k = 0,1,2,3, … x observations; n = database size; and x = number of observations of the haplotype in the database. When α = 0.05, solving for p_0 yields the population frequency of the upper bound of the 95% upper confidence limit. This value is the reported "95% UCI" in YHRD and is the profile probability: the probability of observing the haplotype after adjusting for sampling uncertainty.

7.1.7.1 For instances in which the haplotype has not been observed in the database, the following formula (derived with $x = 0$ in the Clopper and Pearson formula) is used to estimate the upper bound of the 95% confidence limit, p_0:

$$p_0 = 1 - \alpha^{1/n}$$

where α = 0.05 or 5% and n = database size. For example, if a particular haplotype is not observed in a database of 2000 haplotypes, then the 95% upper confidence limit is estimated to be $1 - (0.05)^{1/2000} = 0.0014977$ or 1 in every 778 haplotypes. Note that when α = 0.05, p_0 is very close to 3/n (e.g., 3/2000 = 0.0015).

7.1.8 The likelihood ratio (LR) of the upper bound frequency estimate describes how much more likely are the results if the matching profiles are from the same source or the same paternal lineage rather than if one profile is from an unknown, unrelated individual or lineage. The LR should be calculated:

[11] The Clopper and Pearson formula calculates the exact confidence interval. The exact confidence interval is a cumulative binomial distribution for all values from 0 to x matches given a sample of size *n* and frequency *p*. The listed formula is for the upper limit of a one-tailed confidence interval.

$$\text{LR} = \frac{1}{95\%\ UCI}$$

7.1.8.1 YHRD provides the 95% UCI expressed as 1 in *X*. The LR is the reciprocal of the 95% UCI, which simplifies to LR = *X*.

8 Suggested Reporting Language

The results and/or conclusions for specimens subjected to DNA analysis will generally be reported in narrative form. The formatting and administrative information required in a report are described in the appropriate *FBI Laboratory Operations Manual* practices and the *DNA Procedures Manual*. For guidance on reporting language for Introductory Statements, Amplification Kit Used, Comparisons to Previously Reported Results, Alternate Reference Samples, Elimination Samples, and Differentially Extracted Samples refer to the appropriate interpretation protocols of the *DNA Procedures Manual* (i.e., DNA 233).

8.1 Report Wording Examples

See section 8.3 for endnote language, denoted A-F in these examples.

Direct Comparisons

[Match]

The Y-STR typing results from item 1 were interpreted as originating from two individuals. The major contributor profile from item 1 is 1,100 times more likely if JAMES is the major contributor than if an unknown, unrelated male is the major contributor.[A]

Person of Interest (POI)	Likelihood Ratio (LR)[B]	Level of Support[C]
JAMES	1,100	Moderate support for **Inclusion**

The following individuals are excluded as potential contributors to the Y-STR typing results obtained from item 1:[A]

- JONES
- WHITE

[Uninformative]

The Y-STR typing results from item 1 were interpreted as originating from one individual. The Y-STR typing results from item 1 are equally likely[D] if JAMES is the contributor than if an unknown, unrelated male is the contributor.[A]

Person of Interest (POI)	Likelihood Ratio (LR)[B]	Level of Support[C]
JAMES	1	Uninformative

Lineage Comparisons

[Paternal Relative as Alternate Reference]

Information provided by [person, agency] identifies BROWN as the biological father of JAMES GARCIA. The Y-STR typing results obtained from Item 2 and BROWN are the same; therefore, item 2 could have originated from GARCIA.[A] These results are 310 times more likely if item 2 is from GARCIA than if item 2 is from an unknown, unrelated paternal lineage.

Likelihood Ratio (LR)[B]	Level of Support[C]
310	Moderate support for **Inclusion**

[Paternal Relatedness – Inclusion]

Information provided by [person, agency] identifies JOHNSON as the potential biological brother of MILLER. The Y-STR typing results from JOHNSON and MILLER are the same; therefore, they could be biological brothers.[A] These results are 960 times more likely if JOHNSON and MILLER are paternal relatives than if JOHNSON is from an unknown, unrelated paternal lineage.

Likelihood Ratio (LR)[B]	Level of Support[C]
960	Moderate support for **Inclusion**

[Paternal Relatedness – Exclusion]

Information provided by [person, agency] identifies JOHNSON as the potential biological brother of MILLER. Based on the Y-STR typing results, JOHNSON is excluded as a biological brother of MILLER.[A]

[Paternal Relatedness – Inconclusive]

Information provided by [person, agency] identifies SPARKS as the potential biological brother of MILLER. Based on the Y-STR typing results, no conclusion as to the possible biological relationship between SPARKS and MILLER can be made.[E]

Assumed Contributors (e.g., intimate sample or consensual partner)

[No DNA Unlike]

The Y-STR typing results for item 3 indicate the presence of a single male individual. No Y-STR typing results unlike JONES were obtained from item 3; therefore, no comparisons were made to WHITE.

[Match to Deduced Profile]

The Y-STR typing results from item 3 were interpreted as originating from two individuals, one of whom is JONES. The Y-STR typing results unlike JONES are 230 times more likely if WHITE is the contributor than if an unknown, unrelated male is the contributor.[A]

Person of Interest (POI)	Likelihood Ratio (LR)[B]	Level of Support[C]
WHITE	230	Moderate support for **Inclusion**

BROWN is excluded as a potential contributor to item 3.[A]

Indistinguishable Mixture Comparisons

[Unsub]

The Y-STR typing results for item 1 indicate a mixture of male individuals and are suitable for comparison purposes. Because these mixture results cannot be attributed to individual contributors, they are not suitable for matching purposes; however, they may be used for exclusionary purposes.

[Indistinguishable comparison]

The Y-STR typing results for item 1 indicate the presence of DNA from three male individuals. Because these mixture results cannot be attributed to individual contributors, they are not suitable for matching purposes; however, they may be used for exclusionary purposes. SMITH is excluded as a potential contributor of the DNA obtained from item 1.[A] No conclusion can be provided for JONES.

8.2 Other Reportable DNA Typing Results

8.2.1 When no reference sample is provided for comparison, the results should be reported as follows:

> *"The Y-STR typing results for item 1 indicate a single male individual and are suitable for comparison purposes."*

> *"The Y-STR typing results for item 1 indicate a mixture of male individuals and are suitable for comparison purposes."*

8.2.2 For samples for which insufficient DNA is recovered for DNA typing, this information should be reported as follows:

> *"No Y-STR typing results[F] were obtained from item 1; therefore, no comparisons could be made to SMITH."*

8.2.3 Y-STR typing results may be obtained that are suitable for entry into the Combined DNA Index System (CODIS) or other appropriate database. Refer to the appropriate DNA procedure (i.e., DNA 233) for reporting this information.

8.3 Associated Endnotes for Reporting Language

[A] Barring mutation, any male relative within the same paternal lineage has the same Y-STR profile and would also be expected to be included/excluded as a potential contributor.

[B] The likelihood ratio is a statistical approach that compares the probabilities of observing the DNA results under two alternative propositions. Calculations were performed using the African American, Caucasian, and Hispanic populations in the Y Chromosome Haplotype Reference Database (release xxx). The lowest calculated likelihood ratio is reported.

[C] These likelihood ratio ranges provide the following support for Y-STR conclusions:

Likelihood Ratios:	Qualitative Equivalent:
1	Uninformative
2 to <100	Limited support for **Inclusion**
100 to <10,000	Moderate support for **Inclusion**
10,000 to <1,000,000	Strong support for **Inclusion**

[D] This conclusion is drawn when the likelihood ratio is equal to 1; this comparison is uninformative.

[E] A paternal lineage comparison is declared inconclusive when profiles from putative male relatives from the same paternal lineage are found to differ at only one locus, considering the potential that allele dropout occurred.

[F] Insufficient DNA quality and/or quantity can affect the ability to generate a DNA typing result.

9 Sampling

Not applicable.

10 Measurement Uncertainty

Not applicable.

11 Limitations

11.1 These procedures do not exhaust the possible list of the results that may be encountered by the Examiner. For those results not specifically described, conclusions should be drawn using the procedures given for the results above that are similar in concept and/or origin.

11.2 It is sometimes necessary to consume a sample in its entirety to ensure that the best attempt possible is made to obtain DNA typing results for comparison purposes. Should the total consumption of a sample be required, an Examiner should obtain and record permission from the contributing agency or other responsible office prior to testing.

11.3 A paternal lineage consists of those male relatives to whom the same Y-chromosome has been transmitted from a common ancestor. Barring mutation, all male relatives within the same paternal lineage have the same Y-STR profile. Attribution of the Y-STR typing results to a single individual, to the exclusion of relatives in the paternal lineage, is not possible based on Y-chromosome loci. Additionally, unrelated individuals may exhibit the same Y-STR profile.

12 Safety

Not applicable.

13 References

DNA Procedures Manual

AmpF*l*STR Yfiler PCR Amplification Kit, User's Manual, Applied Biosystems, 2004 Part number 4358101 Rev. A.

Mulero, J.J., Chang, C.W., Calandro, L.M., Green, R.L., Li, Y, Johnson, C.L., and Hennessy, L.K. Development and validation of the AmpF*l*STR YFiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system, *J. Forensic Sci.* (2006) 51:64-75.

Clopper, C., Pearson, E. The use of confidence or fiducial limits illustrated in the case of binomial. *Biometrika* (1934) 26:404-413.
Ge, J., Budowle, B., Planz, J.V., Eisenberg, A.J., Ballantyne, J., and Chakraborty R. US forensic Y-chromosome short tandem repeats database. *Legal Med* (2010) 12:289-295.

Kayser, M., Brauer, S., Willuweit, S., Schadlich, H., Batzer, M.A., Zawacki, J., Prinz, M., Roewer, L., and Stoneking, M. Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations. *J. Forensic Sci.* (2002) 47:513-519.

Butler, J.M., Decker, A.E., Kline, M.C., and Vallone, P.M. Chromosomal duplications along the Y-Chromosome and their potential impact on Y-STR interpretation. *J. Forensic Sci.* (2005) 50:853-859.

Butler, J.M., Schoske, R. Duplication of DYS19 flanking regions in other parts of the Y chromosome. *Int. J. Legal Med* (2004) 118:178-183.

SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Testing Laboratories. Most current version. Available at SWGDAM.org/docs.html.

Rev. #	Issue Date	History
6	10/02/19	Entire document revised and reorganized for clarity. Statistical tool transitioned from US Y-STR database to YHRD. Report language updated to reflect YHRD statistics and to mirror autosomal DNA typing. Moved Y-chromosome figure to Appendix B.
7	12/15/20	6.4.1: added "known" to "reference" 6.6: changed "profiles" to "comparisons" 6.7.1.1, 7.1, 7.1.5, 7.1.6 and renumbered as needed: updated throughout to include calculations for Eskimo Aleut populations when case is submitted from Alaska 8.1: under [Paternal Relative as Alternate Reference] changed "male" to "paternal lineage"

Approval

Redacted - Signatures on File

DNA Technical Leader Date: 12/14/2020

DCU Chief Date: 12/14/2020

Appendix A: *GMIDX Analysis Settings*

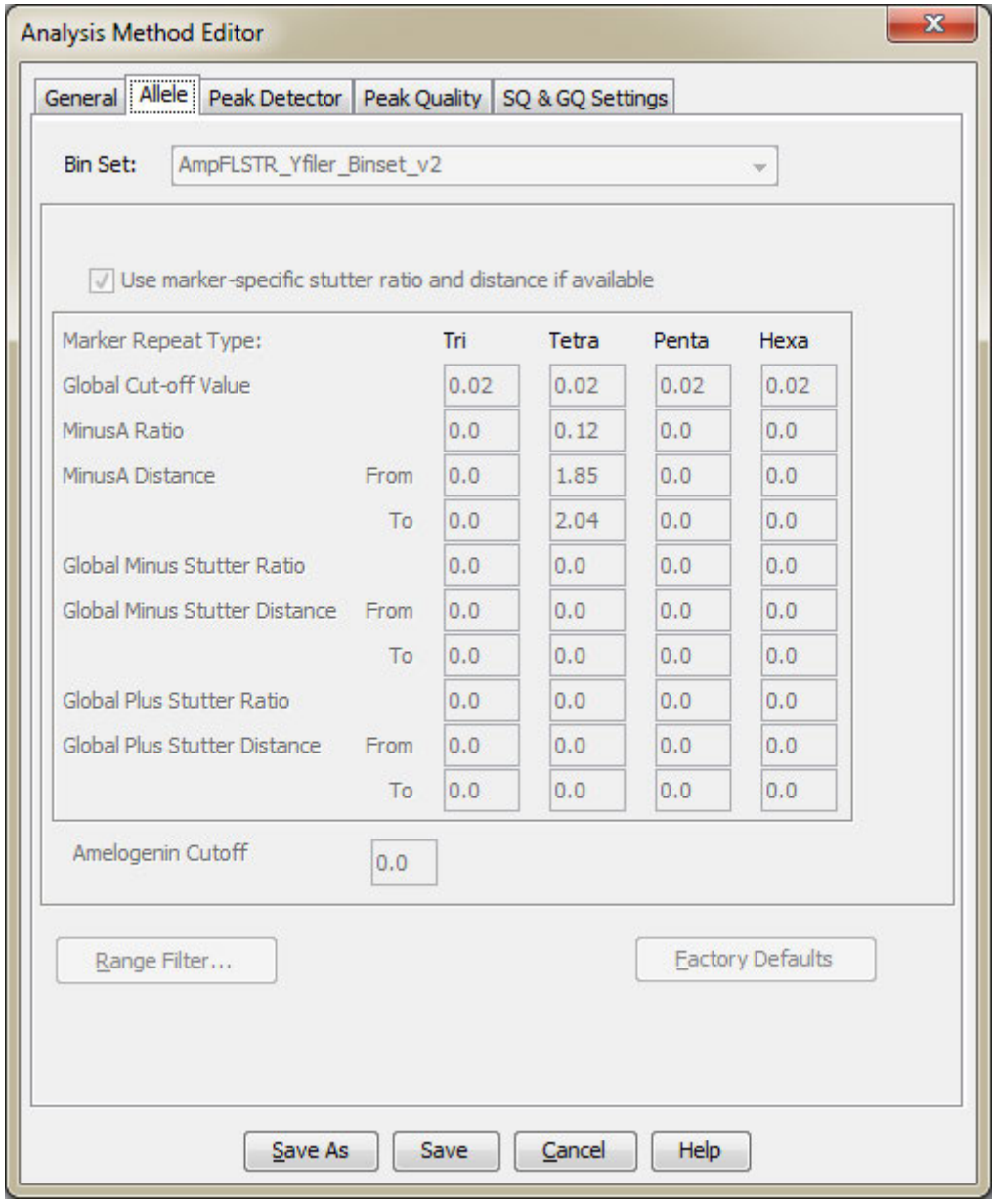

Figure 1 - Allele Tab for YFiler™

Appendix A: *GMIDX Analysis Settings (cont.)*

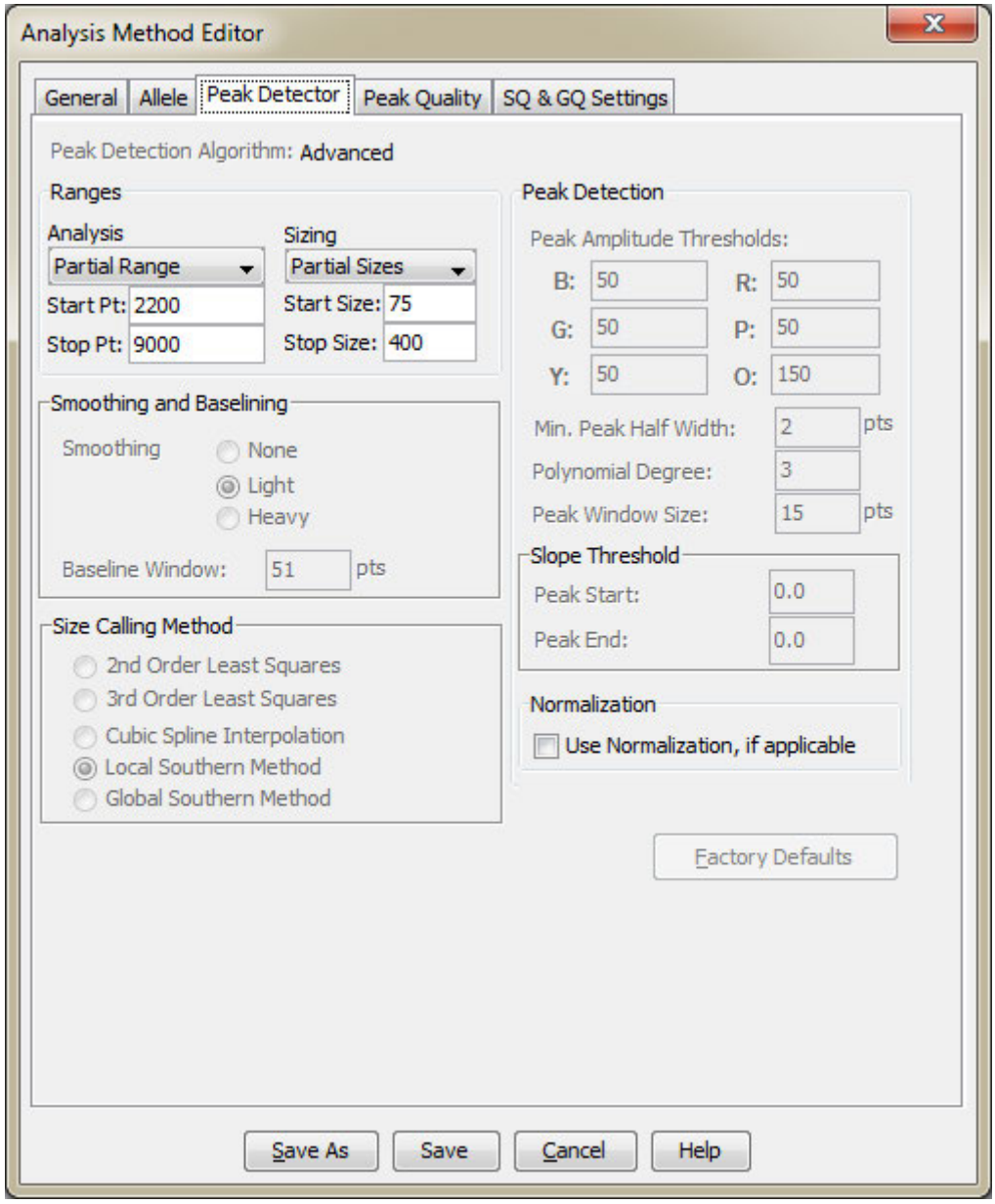

Figure 2 - Peak Detector Tab for YFiler™

Appendix B: Map of the Y-Chromosome

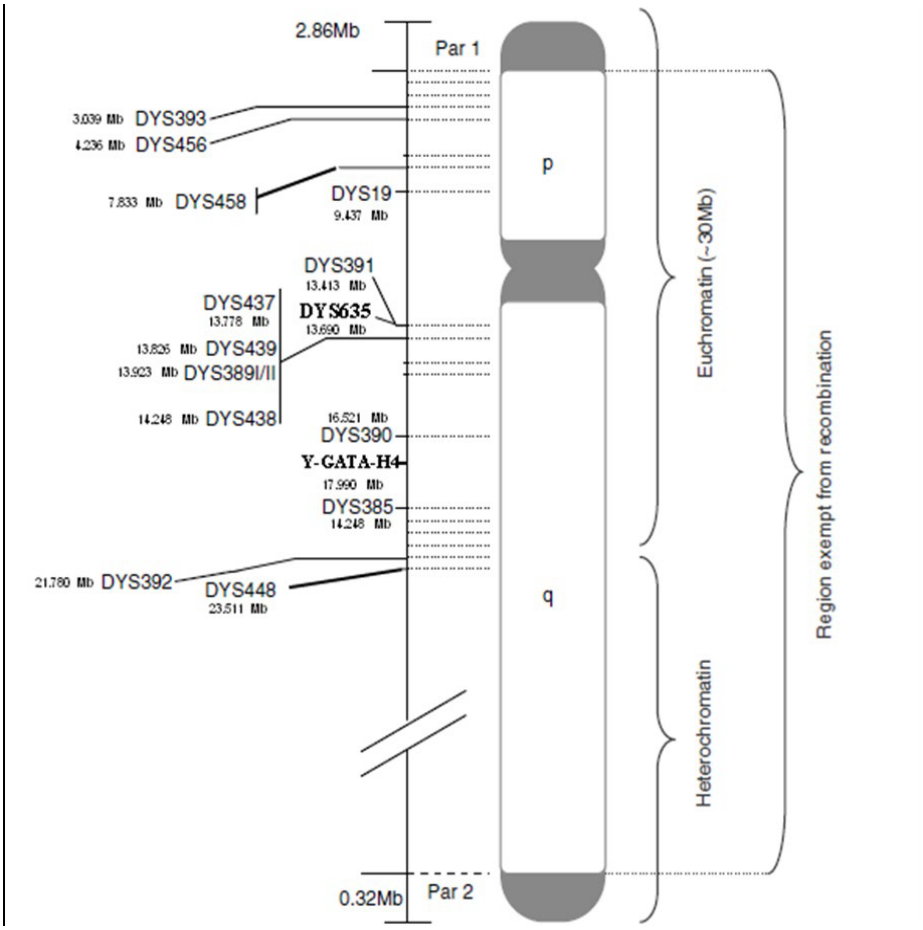

Figure 3 – Y-Chromosome with Yfiler™ loci and their relative locations and distances[12]

[12] Buckleton, J., C. Triggs and S. Walsh. Forensic DNA Evidence Interpretation. CRC Press; Boca Raton, Fl, 2005.